A Unified Cosolvency Model for Calculating Solute Solubility in Mixed Solvents

Abolghasem JOYBAN,* a Nora Yat Knork CHEW,b Hak-Kim CHAN,b Mohammad SABOUR,c and William Eugene ACREE, Jr.d

a School of Pharmacy and Drug Applied Research Center, Tabriz University of Medical Sciences; Tabriz 51664, Iran; b Faculty of Pharmacy, The University of Sydney; Sydney NSW 2006, Australia; c Kimia Research Institute; Tabriz, P.O. Box 51665–171, Iran; and d Department of Chemistry, University of North Texas; Denton 76203–5070, U.S.A.

Received January 18, 2005; accepted March 10, 2005

Organic solvents are amongst the most powerful solubilization agents for a large number of water-insoluble drugs. A number of equations has been reported for mathematical representation of solute solubility in mixed solvents. The question is then posed—is there a mathematical difference between these models? To address this point, it has been demonstrated that all cosolvency models could be made equivalent by using algebraic manipulations. In order to familiarize the readers with the available cosolvency models, they are briefly reviewed. The models can be divided into two mathematical categories, i.e. linear and non-linear models. The linear models include: the log-linear, extended Hildebrand solubility approach, excess free energy equations, combined nearly ideal binary solvent/Redlich-Kister equation and Margule equations which can be converted to a general single model which expresses the logarithm of mole fraction solubility of a solute as a power series of volume fraction of the cosolvent. The non-linear models include the mixture response surface methods, two step solvation model and modified Wilson model which can be converted to a non-linear general form. Also, it has been shown that both the general single model and a non-linear general model are mathematically identical. To show the applicability of the models on real experimental data, 35 data sets have been collected from the literature. Both linear and nonlinear models produced comparable accuracies when an equal number of constant terms was employed in numerical analyses.

Key words solubility; solvent effect; mixed solvent; mathematical model; cosolvency

Aqueous solubility of drugs is one of the key factors in developing a new drug and the blending of different solvents is a common method to increase the solubility. Apart from experimental determinations of a solute solubility in water-co-solvent mixtures, many mathematical models have been established to describe solute solubility in mixed solvents.1 9) Some of these models are theoretical, while others are semi-theoretical or empirical. While the empirical ones are mainly used to correlate between experimental solubilities and independent variables such as the volume fraction of the cosolvent, the theoretical ones can improve the understanding of solubility behaviour for drugs in mixed solvents.

It has been found that the solute solubility in mixed solvents can be mathematically represented by a single equation. There, however, is a number of equations that can be considered which usually produce comparable results. The question is then posed—is there a mathematical difference between these models? To address this point, it has been demonstrated in this work that all the suggested cosolvency models could be made equivalent by using algebraic manipulations. Based on these manipulations, a unified cosolvency model has been proposed in the present study.

Theoretical Treatment

The log-linear relationship,20 extended Hildebrand solubility approach,14 excess free energy equations,21) the simplest form of the mixture response surface method, and the combined nearly ideal binary solvent/Redlich-Kister (CNIBS-RK) model21) have been converted to a general single model, GSM.21) GSM correlates the logarithm of a solute solubility as a polynomial function of cosolvent’s volume fraction as:

\[ \ln X_n = M_0 + M_1 f_1 + M_2 f_1^2 + M_3 f_1^3 + \cdots \]  

(1)

Where \( X_n \) is the mole fraction solubility of the solute, \( f_1 \) is volume fraction of cosolvent in the absence of the solute and \( M_0 - M_4 \) are the model constants. Before the unified cosolvency model derived in this study is discussed, different non-linear mathematical models on solubility was first reviewed.

Mixture Response Surface Model Statistically based mixture response surface methods, MRS,14) have also been proposed for correlative purposes and these models are as follows:

\[ \ln X_n = \beta_0 f_1^{\prime} + \beta_1 f_1^{\prime} f_2^{\prime} + \beta_2 \left( \frac{1}{f_2} \right) + \beta_3 \left( \frac{1}{f_2} \right) + \beta_4 \left( \frac{1}{f_2} \right) + \beta_5 \left( \frac{1}{f_2} \right) \]

(2a)

\[ \ln X_n = \beta_0 f_1^{\prime} + \beta_1 f_1^{\prime} f_2^{\prime} + \beta_2 \left( \frac{1}{f_2} \right) + \beta_3 f_1^{\prime} f_2^{\prime} \]

(2b)

\[ \ln X_n = \beta_0 f_1^{\prime} + \beta_1 f_1^{\prime} f_2^{\prime} + \beta_2 \left( \frac{1}{f_2} \right) + \beta_3 \left( \frac{1}{f_2} \right) + \beta_4 \left( \frac{1}{f_2} \right) + \beta_5 f_1^{\prime} f_2^{\prime} \]

(2c)

In which \( \beta_0 - \beta_5 \) are the model’s parameters and \( f_1^{\prime} \) and \( f_2^{\prime} \) are given by \( f_1^{\prime} = 0.96 f_1 + 0.02 \) and \( f_2^{\prime} = 0.96 f_2 + 0.02 \) in which \( f_1 \) is volume fraction of water.4) Modified Wilson Model The modified Wilson model (MWM), is another possibility which is shown below:

\[ \ln \left( \frac{X_n}{X_0} \right) = - \frac{f_1 \left[ 1 - \ln \left( \frac{X_n}{X_0} \right) \right]}{f_1 + f_2 A_{12}^{w}} - \frac{f_2 \left[ 1 - \ln \left( \frac{X_n}{X_0} \right) \right]}{f_1 + f_2 A_{12}^{w}} \]

(3a)

Where \( X_0 \) and \( X_n \) denote the mole fraction solubility in neat cosolvent and water, respectively.3) It was shown that a simplified form of the modified Wilson model, SMW,9) is able to calculate solute solubility in water–cosolvent mixtures more accurate than MWM, although this simplification is not successful in the case of solubility prediction in non-aqueous binary solvents.3) Thus the SMW is:

\[ - \ln X_n = - \frac{f_1 \left( 1 + \ln X_1 \right)}{f_1 + f_2 A_{12}^{w}} - \frac{f_2 \left( 1 + \ln X_2 \right)}{f_1 + f_2 A_{21}^{w}} \]

(3b)

Where \( A_{12}^{w}, A_{21}^{w}, A_{12}^{w} \) and \( A_{21}^{w} \) are adjustable parameters of the models which can be evaluated via a nonlinear least squares analysis.
Phenomenological Model Khosravi and Connors\(^7\) developed a phenomenological model for describing the solvent effects on the equilibrium solubility of a solute in a binary solvent mixture. The model could be represented as:

\[-kT \ln X_m = kT \ln X_2 + \frac{a\beta f_1 f_2 + b\beta f_2^2}{f_1^2 + \beta f_2 f_1} \]

(4)

Where \(k\) is the Boltzman's constant, \(T\) is the absolute temperature, and \(a\), \(b\), \(\beta\), and \(f_i\) are the model constants.\(^7\)

Unified Cosolvency Model It can be expected that for a given phenomenon, a single model should be able to mathematically represent the experimental solubility data. However, as discussed above, there have been many different equations. Each of them has different errors in the results when matched against a training data set, due to the different assumptions and simplifications employed.

Substitution of \(f_i\) with \((1-f_i)\) in non-linear Eqs. 2b, 2c and 3a, 3b with subsequent rearrangements yields:

\[\ln X_m = \frac{J_{2b} + J_{2c} f_2^2 + J_{3a} f_3^2 + \ldots}{K_{2b} + K_{2c} f_2 + K_{3a} f_3 + \ldots} \]

(5a)

Where \(J_{2b}, J_{2c}, \text{ and } K_{2b}, K_{2c}, K_{3a}\) are the model constants computed by using a non-linear least square analysis. Since the \(X_m\) terms on the left-hand side of Eqs. 1 and 5a are the same, it is possible to write:

\[M_m + M_{f1} f_1 + M_{f2} f_2 + M_{f3} f_3 + \ldots = \frac{J_{2b} + J_{2c} f_2^2 + J_{3a} f_3^2 + \ldots}{K_{2b} + K_{2c} f_2 + K_{3a} f_3 + \ldots} \]

(5b)

By multiplying \((M_m + M_{f1} f_1 + M_{f2} f_2 + M_{f3} f_3 + \ldots)\) in \((K_{2b} + K_{2c} f_2 + K_{3a} f_3 + \ldots)\) for Eq. 5b and subsequent rearranging, it is possible to obtain:

\[K_{2b} M_m + K_{2b} M_{f1} f_1 + K_{2c} M_{f2} f_2 + K_{3a} M_{f3} f_3 + \ldots + J_{2b} f_2 f_1 + J_{2c} f_2 f_2 + J_{3a} f_3 f_3 + \ldots + K_{2b} M_m f_2 + K_{2b} M_{f1} f_1 f_2 + K_{2c} M_{f2} f_2 f_2 + K_{3a} M_{f3} f_3 f_2 + \ldots + J_{2b} f_2^2 f_1 + J_{2c} f_2^2 f_2 + J_{3a} f_3^2 f_3 + \ldots\]

(5c)

Further rearrangement of Eq. 5c can produce:

\[K_{2b} M_m f_2 + K_{2b} M_{f1} f_2 f_1 + K_{2c} M_{f2} f_2 f_2 + K_{3a} M_{f3} f_2 f_3 + \ldots + J_{2b} f_2^2 f_1 + J_{2c} f_2^2 f_2 + J_{3a} f_3^2 f_3 + \ldots\]

(5d)

Since \(K_{2b}, K_{2c}, \text{ and } K_{3a}\) and other terms in parentheses are constant values for a given binary system, it is possible to re-write Eq. 5d as Eq. 5e.

\[A_m + A_{f1} f_1 + A_{f2} f_2 + A_{f3} f_3 + \ldots = J_{2b} f_2 f_1 + J_{2c} f_2 f_2 + J_{3a} f_3 f_3 + \ldots\]

(5e)

As an example, Eq. 3b could be rewritten as:

\[-\ln X_m = \ln (f_1 + f_2 \lambda_{2a}) + \frac{\ln (1 + \ln X_1)}{f_1 \lambda_{2a}^2 + f_2} \]

(3c)

By replacing \(f_2\) with \((1-f_2)\), \((1+\ln X_1)\) and \((1+\ln X_2)\) with \(\lambda_2\) and \(\lambda_4\) and subsequent rearranging the following equation can be obtained:

\[-\ln X_m = -\frac{f_2 \lambda_2}{f_1 \lambda_2^2 + f_2} + \frac{(1-f_2) \lambda_4}{f_1 \lambda_4^2 + 1 - f_1} \]

(3d)

or:

\[-\ln X_m = -\frac{f_2 \lambda_2 (f_1 \lambda_2^2 + 1 - f_1)}{(f_2 \lambda_2^2 + 1 - f_1) (f_1 \lambda_4^2 + 1 - f_1)} \]

(3e)

or:

\[-\ln X_m = -\frac{f_2 \lambda_2 (f_1 \lambda_2^2 + 1 - f_1)}{(f_2 \lambda_2^2 + 1 - f_1) K_1 (f_1 \lambda_4^2 + 1 - f_1)} \]

(3f)

or:

\[-\ln X_m = -\frac{f_2 \lambda_2 (f_1 \lambda_2^2 + 1 - f_1)}{(f_2 \lambda_2^2 + 1 - f_1) (f_1 \lambda_4^2 + 1 - f_1)} \]

(3g)

Since \(X_m\) in Eq. 3g are constant for a given solute in a binary solvent system, it is possible to summarize Eq. 3g as:

\[\ln X_m = \frac{J_{2b} + J_{2c} f_2 + J_{3a} f_3^2}{K_{2b} + K_{2c} f_2 + K_{3a} f_3^2} \]

(3h)

Where \(J_{2b} = (-\lambda_{2a}^2 + \lambda_{4a}^2), J_{2c} = (-\lambda_{2a}^2 + \lambda_{4a}^2 + 2\lambda_{2a}^2 + \lambda_2 - \lambda_{4a}^2), J_{3a} = (1 - \lambda_{2a}^2 + \lambda_{4a}^2 + 2\lambda_{2a}^2 + \lambda_2 - \lambda_{4a}^2), \lambda_2 = \lambda_{2a}, \lambda_4 = \lambda_{4a}\) and \(K_{2b} = (-\lambda_{2a}^2 + \lambda_{4a}^2 + 2\lambda_{2a}^2 + \lambda_2 - \lambda_{4a}^2), K_{2c} = (1 - \lambda_{2a}^2 + \lambda_{4a}^2 + 2\lambda_{2a}^2 + \lambda_2 - \lambda_{4a}^2)\).

From these equations, we can summarise all cosolvency models as a power series of volume fraction of the cosolvent, GSM model. The GSM was used in earlier works by Martin and co-workers\(^10\) and a mathematical justification for GSM has been provided.\(^8\) The above mentioned mathematical manipulations showed that the non-linear cosolvency models could also be converted to GSM. These findings are not however unexpected as it is generally the case that a definite experimental phenomenon, like drug solubility in water-cosolvent mixtures, would have a single mathematical representation. Here it has been shown that this is in fact true for the cosolvency models. However, the accuracy of these models differs from each other. This is because the models employed a different arrangement of independent variables.

Computational Methods To assess the accuracy of the equations, the experimental \(X_m\) values were fitted into the equations and the mean percentage deviation (MPD) between experimental and calculated \(X_m\) values was considered as an accuracy criterion. The MPD is defined as:

\[\text{MPD} = \frac{100}{N} \sum \frac{|\text{calculated} - \text{observed}|}{\text{observed}} \]

where \(N\) is the number of experimental data points in each set. The mean value of MPDs is denoted as overall MPD (OMPD) and is given by:

\[\frac{\sum_{i=1}^{N} \text{MPD}}{N} \]

The computations could be carried out using various statistical softwares. As an example, a computer program using the SPSS was presented in the appendix section. The program calculates various statistical data of the most comprehensive equation including model constants and the MPD value for the solubility of oxolinic acid in water + ethanol mixture.

Results and Discussion In the present study, we tested the accuracy of the equations by fitting the experimental data sets (for details see Table 1) to the equations and considered the number of constant terms, the MPD and OMPD values. The differences between the OMPD values for all of the models discussed above using the equal number of constant terms, i.e. 4—6, were evaluated using ANOVA and the mean differences (details were not shown here) in all cases were statistically significant (ANOVA, \(p<0.0005\)). This finding is in agreement with a previously reported result.\(^12\) The differences in OMPDs produced by the equations could be justified by considering the different assumptions used to derive the models, the simplifications made during the model development process, different independent variables and the numerical analysis method. As an example, it has been shown that two different numerical methods in obtaining the model constants of the CNIBS/R-K model produced various MPD values.\(^13\)

As shown in the theoretical treatment, Eq. 3b could be made equivalent to Eq. 3h, and by doing this the OMPDs and standard deviations for Eqs. 3b and 3h were 10.2±7.0 and 10.1±15.2, respectively. The OMPD difference is insignificant (paired t-test, \(p>0.99\)). The OMPD obtained from Eqs.
Table 1. Details of Solubility Data, the Number of Data Points in Each Set (N), the Mean Percentage Deviation (MPD) of Eq. 6 and the References

<table>
<thead>
<tr>
<th>No.</th>
<th>Cosolvent</th>
<th>Solute</th>
<th>N</th>
<th>MPD</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Acetonitrile</td>
<td>Theophylline</td>
<td>17</td>
<td>3.1</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Dimethylformamide</td>
<td>Caffeine</td>
<td>11</td>
<td>4.0</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>Dimethylformamide</td>
<td>Sulphadiazine</td>
<td>14</td>
<td>2.2</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Dioxane</td>
<td>Caffeine</td>
<td>16</td>
<td>3.7</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Dioxane</td>
<td>p-Hydroxybenzoic acid</td>
<td>13</td>
<td>5.0</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>Dioxane</td>
<td>Paracetamol</td>
<td>17</td>
<td>7.7</td>
<td>24</td>
</tr>
<tr>
<td>7</td>
<td>Dioxane</td>
<td>Phenacetin</td>
<td>13</td>
<td>3.7</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>Dioxane</td>
<td>Sulphadiazine</td>
<td>17</td>
<td>11.3</td>
<td>26</td>
</tr>
<tr>
<td>9</td>
<td>Dioxane</td>
<td>Sulphadimidine</td>
<td>19</td>
<td>14.6</td>
<td>26</td>
</tr>
<tr>
<td>10</td>
<td>Dioxane</td>
<td>Sulphanethiazolate</td>
<td>19</td>
<td>17.0</td>
<td>27</td>
</tr>
<tr>
<td>11</td>
<td>Dioxane</td>
<td>Sulphanethoxazolone</td>
<td>15</td>
<td>9.7</td>
<td>26</td>
</tr>
<tr>
<td>12</td>
<td>Dioxane</td>
<td>Sulphapyridine</td>
<td>17</td>
<td>10.2</td>
<td>28</td>
</tr>
<tr>
<td>13</td>
<td>Dioxane</td>
<td>Sulphanethoxypridazine</td>
<td>18</td>
<td>8.7</td>
<td>26</td>
</tr>
<tr>
<td>14</td>
<td>Dioxane</td>
<td>Sulphanilamide</td>
<td>16</td>
<td>11.3</td>
<td>29</td>
</tr>
<tr>
<td>15</td>
<td>Dioxane</td>
<td>Sulphasomidine</td>
<td>21</td>
<td>12.5</td>
<td>30</td>
</tr>
<tr>
<td>16</td>
<td>Dioxane</td>
<td>Theobromine</td>
<td>11</td>
<td>1.8</td>
<td>30</td>
</tr>
<tr>
<td>17</td>
<td>Dioxane</td>
<td>Theophylline</td>
<td>21</td>
<td>5.9</td>
<td>31</td>
</tr>
<tr>
<td>18</td>
<td>Ethanol</td>
<td>Furosemide</td>
<td>13</td>
<td>6.2</td>
<td>32</td>
</tr>
<tr>
<td>19</td>
<td>Ethanol</td>
<td>Oxolinic acid</td>
<td>11</td>
<td>0.8</td>
<td>33</td>
</tr>
<tr>
<td>20</td>
<td>Ethanol</td>
<td>Paracetamol</td>
<td>13</td>
<td>6.6</td>
<td>24</td>
</tr>
<tr>
<td>21</td>
<td>Ethanol</td>
<td>Sulphanethazine</td>
<td>11</td>
<td>4.9</td>
<td>34</td>
</tr>
<tr>
<td>22</td>
<td>Ethanol</td>
<td>Sulphanilamide</td>
<td>12</td>
<td>3.0</td>
<td>34</td>
</tr>
<tr>
<td>23</td>
<td>Ethylene glycol</td>
<td>Naphthalene</td>
<td>18</td>
<td>1.4</td>
<td>7</td>
</tr>
<tr>
<td>24</td>
<td>Ethylene glycol</td>
<td>Theophylline</td>
<td>17</td>
<td>1.9</td>
<td>7</td>
</tr>
<tr>
<td>25</td>
<td>Glycerine</td>
<td>Furosemide</td>
<td>12</td>
<td>3.3</td>
<td>32</td>
</tr>
<tr>
<td>26</td>
<td>Methanol</td>
<td>Theophylline</td>
<td>13</td>
<td>1.4</td>
<td>7</td>
</tr>
<tr>
<td>27</td>
<td>Propylene glycol</td>
<td>Butyl p-aminobenzoate</td>
<td>11</td>
<td>4.7</td>
<td>35</td>
</tr>
<tr>
<td>28</td>
<td>Propylene glycol</td>
<td>Butyl p-hydroxybenzoate</td>
<td>11</td>
<td>4.0</td>
<td>35</td>
</tr>
<tr>
<td>29</td>
<td>Propylene glycol</td>
<td>Ethyl p-aminobenzoate</td>
<td>11</td>
<td>3.0</td>
<td>35</td>
</tr>
<tr>
<td>30</td>
<td>Propylene glycol</td>
<td>Ethyl p-hydroxybenzoate</td>
<td>11</td>
<td>5.1</td>
<td>35</td>
</tr>
<tr>
<td>31</td>
<td>Propylene glycol</td>
<td>Furosemide</td>
<td>13</td>
<td>5.3</td>
<td>32</td>
</tr>
<tr>
<td>32</td>
<td>Propylene glycol</td>
<td>Methyl p-aminobenzoate</td>
<td>11</td>
<td>2.0</td>
<td>35</td>
</tr>
<tr>
<td>33</td>
<td>Propylene glycol</td>
<td>Methyl p-hydroxybenzoate</td>
<td>11</td>
<td>3.2</td>
<td>35</td>
</tr>
<tr>
<td>34</td>
<td>Propylene glycol</td>
<td>Propyl p-aminobenzoate</td>
<td>11</td>
<td>3.7</td>
<td>35</td>
</tr>
<tr>
<td>35</td>
<td>Propylene glycol</td>
<td>Propyl p-hydroxybenzoate</td>
<td>11</td>
<td>4.4</td>
<td>35</td>
</tr>
</tbody>
</table>

Fig. 1. Experimental Solubility of Oxolinic Acid in Water–Ethanol Mixtures, and the Reproduced Curves Using Eqs. 1, 3h and 6

1 and 3h were 9.0±5.4 and 10.2±7.0, respectively, and again the mean difference is not significant (paired t-test, p>0.57). Figure 1 shows the reproduced solubility profile of oxolinic acid at different ethanol concentrations using the different models discussed. The high standard deviation for Eq. 3h is related to the nature of the iteration method, where standard errors of all model constants for non-linear models are also high. As an example, J0 (of Eq. 3h) for data of theophylline in water-acetonitrile mixtures is −0.14 and its standard error is 14265.50. To reduce the standard error of the model constants for non-linear equations, it is suggested to employ more data points. This, however, is not a suitable solution, when the aim of a research is to optimise solvent composition of a binary solvent mixture for solubilization and/or desolubilization purposes, and to give a fast and low-costly method. But we should be careful not to include too many experimental data points since the purpose of mathematical modeling (i.e. prediction) will be lost. Previously we have used trained mathematical models by five experimental data, which provides accurate predictions.

In conclusion, it has been shown that all cosolvency models from the literature could be made mathematically equivalent and with different cosolvency models described above regarding solute solubility data in a binary solvent mixture, researchers have the dilemma of having many models to choose from in their practical applications. From the work carried out both in this paper and in previous studies by the group, Eq. 6 is recommended for practical applications. Equation 6 is:

\[
\ln X_a = \ln X_0 + \frac{Q}{1+\ln \frac{Q(L_1 \ln f_1)}{1-Q}}
\]

where \(Q\) is the model constant and \(q\) is usually 2—3. As has
been shown in Table 1, Eq. 6 provides the most accurate calculations and its main advantages over the others are:
- Simple and reliable calculations (see the Appendix)
- Capability of calculating the solute solubility in mixed solvents at different temperatures
- Capability of describing multiple solubility maxima in mixed solvents
- Capability of calculating the solubility of structurally related drugs in mixed solvents
- Representation of the solubility of polymorphs in mixed solvents
- Possibility of extending its solubility prediction capabilities to ternary solvents using sub-binary data
- Possessing as many curve-fitting parameters as needed for accurate representation of experimental data in mixed solvents
- Capability of describing other physico-chemical properties of solutes in mixed solvent systems.

If this equation is used, then the pharmaceutical chemist is likely to be able to reduce the length of operation of the drug solubilization/desolubilization process using solvent mixtures.

Acknowledgements The authors would like to thank the Australian Department of Education, Training and Youth Affairs and the University of Sydney for providing the IPRS and IPA scholarships.

References

Appendix
A computer program using the SPSS software for calculation of solubility of a solute in binary aqueous cosolvent mixture using Eq. 6 (with \( q = 2 \)).

* F1: Volume fraction of the cosolvent.
* LXM: Logarithm of mole fraction solubility of the solute in mixed solvent.

DATA LIST FREE/F1 LXM.
BEGIN DATA.
0.00 -13.74
0.10 -12.74
0.20 -12.13
0.30 -11.81
0.40 -11.52
0.50 -11.28
0.60 -11.02
0.70 -10.81
0.80 -10.77
0.90 -10.99
1.00 -11.71
END DATA.

TITLE 'Solubility of oxolinic acid in water-ethanol at 25 °C'.
* LX1: Logarithm of mole fraction solubility of the solute in neat cosolvent.
* LX2: Logarithm of mole fraction solubility of the solute in neat water.
COMPUTE LX1 = -11.71.
COMPUTE LX2 = -13.74.
* F2: Volume fraction of water.
COMPUTE F2 = 1-F1.
COMPUTE Q0 = F1*F2.
COMPUTE Q1 = F1*F2*(F1-F2).
COMPUTE Q2 = F1*F2*(F1-F2)*(F1-F2).
COMPUTE Y = LXM*F1*LX1*F2*LX2.
REGRESSION /ORIGIN /DEPENDENT Y /METHOD=ENTER Q0 Q1 /
SAVE PRED.
COMPUTE LXM=PRE_1+F1*LX1+F2*LX2.
COMPUTE XMP = EXP(LXM).
COMPUTE XM = EXP(LX2).
COMPUTE MDSP = ABS(100*(XMP-XM)/XM).
DESCRIPTIVES VARIABLES=MDSP /STATISTICS=MEAN.